Elly anak Stephen


# Software Metric for Assessing the Quality of Software Requirements

Master of Science

Elly anak Stephen

Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
2020

2020

# Software Metric for Assessing the Quality of Software Requirements

Elly anak Stephen

A thesis submitted

In fulfillment of the requirements for the degree of Master of Science

(Software Engineering)

Faculty of Computer Science and Information Technology
UNIVERSITI MALAYSIA SARAWAK
2020

## DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. Except where due acknowledgements have been made, the work is that of the author alone. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

……

Signature

Name: Elly anak Stephen

Matric No.: 15020302

Faculty of Computer Science and Information Technology

Universiti Malaysia Sarawak

Date :

## ACKNOWLEDGEMENT

I would like to thank everyone that has given their contribution toward the success of my thesis. I would like to express my gratitude towards my supervisor AP. Dr. Edwin Mit for his patient, advice and guidance throughout this research.

In addition, I would like to thank both of my parents for giving me a chance and encouragement towards the completion of this research.

Furthermore, I would like to thank Universiti Malaysia Sarawak (UNIMAS) for the opportunity in giving me the job as Research Assistant and grant under FRGS/2/2014/ICT07/UNIMAS/02/1.

.

## ABSTRACT

SRS document composes of FR and NFR. However, due to the heterogeneous domain development environment, the quality of the produced SRS document is in question. Software development depends on the quality of the SRS document. Poor quality of the SRS document results in poor software production. A study on the sample of the SRS document shows a lack of standardization of the document structure as well as its FR. A recent researcher manually evaluates the SRS document which is time consuming. Furthermore, the unstructured way in writing FR leads to ambiguity. A recent researcher shows that the ambiguity of natural language can be reduced by restricting the term used in writing FR. Numbers properties for each quality had been introduced to evaluate the written FR. Two methods are adopted in order to standardize the structural and FR. The first method is by comparison with the IEEE 830 ToC. The topic from the ToC is extracted and used to assess the structure of the document. The second method is the RB. It is used to standardize the way of defining the FR. The structural and FR are assessed based on four quality properties which are completeness, consistency, correctness and preciseness. The completeness quality is intended to assess the SRS structure meanwhile the others are for FR. A framework is designed based on quality properties to be assessed. The quality properties are formalized by proposing the equation for measurement. The framework is implemented and the prototype is tested to evaluate its capability. Case study is applied to the prototype. This resulted in two outputs. The first output is by comparison against the IEEE 830 which results in the percentage similarity of structure. Meanwhile, second output resulted from prototype analyst the user input. The overall measurement of the case study is produced by a prototype in the degree of percentage by adding both from the first and second outputs. From the results, the developer can evaluate the maturity of the SRS document. The resulting

measurement of the SRS structure allows the developer to reorganize the document to increase readability. The result from the measurement of FR allows the developer to ensure the consistency, validating and avoiding unspecific usage of the term used in writing the FR.

**Keywords:** Quantitative measurement, software quality, software requirement specification

***Metrik Perisian untuk Menilai Kualiti Keperluan Perisian***

## *ABSTRAK*

*Dokumen SRS mengandungi FR dan NFR. Walau bagaimanapun, disebabkan oleh persekitaran pembangunan domain yang heterogen, kualiti dokumen SRS yang dihasilkan agak meragukan. Pembangunan perisian bergantung pada kualiti dokumen SRS. Lemahnya kualiti dokumen SRS akan menghasilkan pengeluaran perisian yang lemah. Kajian sampel dokumen SRS menunjukkan kekurangan standardisasi struktur dokumen serta FR. Penyelidik baru-baru ini secara manual menilai struktur dokumen SRS yang memakan masa. Selain itu, cara yang tidak tersusun dalam penulisan FR membawa kepada ketidakpastian. Seorang penyelidik baru-baru ini menunjukkan bahawa ketidakpastian dalam bahasa tabii dapat dikurangkan dengan menyekat istilah yang digunakan dalam penulisan FR. Ciri-ciri nombor bagi setiap kualiti telah diperkenalkan untuk menilai FR yang ditulis. Dua kaedah digunakan untuk menyeragamkan struktur dan FR. Kaedah pertama adalah dengan perbandingan dengan IEEE 830 ToC. Topik dari ToC diekstraksi dan digunakan untuk menilai struktur dokumen. Kaedah kedua ialah RB. Ia digunakan untuk menyeragamkan cara mendefinisikan FR. Struktur dan FR dinilai berdasarkan empat sifat kualiti iaitu kelengkapan, konsistensi, kebenaran dan ketepatan. Kualiti kelengkapan ini bertujuan untuk menilai struktur SRS sementara yang lain adalah untuk FR. Rangka kerja direka berdasarkan sifat-sifat kualiti yang akan dinilai. Ciri-ciri kualiti diformalkan dengan mencadangkan persamaan untuk pengukuran. Rangka kerja dilaksanakan dan prototaip diuji untuk menilai keupayaannya. Kajian kes diaplikasikan ke prototaip. Dua hasil dikeluarkan. Keluaran pertama adalah hasil persamaan perbandingan struktur berbanding dengan IEEE 830. Sementara keluaran kedua dihasilkan daripada penganalisis prototaip input pengguna. Pengukuran keseluruhan kajian kes dihasilkan oleh prototaip dalam tahap*

*peratusan hasil pemenambahkan dari hasil pertama dan kedua. Daripada keputusan, pemaju boleh menilai kematangan dokumen SRS. Pengukuran hasil struktur SRS membenarkan pemaju menyusun semula dokumen itu untuk meningkatkan kebolehbacaan. Hasil dari pengukuran FR membolehkan pemaju memastikan konsistensi, mengesahkan dan mengelakkan penggunaan istilah yang tidak spesifik digunakan dalam penulisan.*

***Kata kunci:*** *Kualiti perisian, spesifikasi keperluan perisian, pengukuran kuantitatif*

## TABLE OF CONTENTS


| | | Page |
|---|---|---|
| DECLARATION | | i |
| ACKNOWLEDGEMENT | | ii |
| ABSTRACT | | iii |
| ABSTRAK | | v |
| TABLE OF CONTENTS | | vii |
| LIST OF TABLES | | xi |
| LIST OF FIGURES | | xii |
| LIST OF ABBREVIATIONS | | xiv |
| **CHAPTER 1: INTRODUCTION** | | 1 |
| 1.1 | Background | 1 |
| 1.2 | Problem Statement | 2 |
| 1.3 | Objectives | 4 |
| 1.4 | Hypothesis | 4 |
| 1.5 | Scope of Project | 5 |
| 1.6 | Thesis Structure | 6 |
| 1.7 | Significant of Project | 7 |
| 1.8 | Summary | 8 |
| **CHAPTER 2: LITERATURE REVIEW** | | 9 |

2.1 Introduction 9

2.2 Software Quality 9

2.3 Software Requirement Specification 12

2.3.1 Structural 15

2.3.2 Functional Requirement 17

2.3.3 Natural Language Processing for Requirement Analysis 23

2.3.4 Challenges of Current Approach 24

2.4 Summary 27

**CHAPTER 3: RESEARCH METHODOLOGY** 28

3.1 Introduction 28

3.2 Phase 1: Review on problem and solutions 28

3.3 Phase 2: Defining Problem 29

3.3.1 Propose Rule and Equation 35

3.4 Phase 3: Developing Prototype 44

3.5 Phase 4: Evaluation of Formalize Problem 47

3.6 Summary 48

**CHAPTER 4: IMPLEMENTATION** 50

4.1 Introduction 50

4.2 Prototype Architecture 50

4.2.1 Activity Diagram 51

4.3 Prototype Design 54

4.4 Prototype Sample Output 64

4.5 Summary 81

**CHAPTER 5: DISCUSSION** 82

5.1 Introduction 82

5.2 Case-control Study 82

5.2.1 Structural Assessment 83

5.2.2 Functional Requirement Assessment 86

5.3 Actual Case Study 98

5.3.1 Structural Assessment 99

5.3.2 Functional Requirement Assessment 101

5.3.3 Overall Measurement 108

5.3.4 Discussion 110

5.4 Summary 114

**CHAPTER 6: CONCLUSION** 115

6.1 Introduction 115

6.2 Achievement 115

6.3 Contribution 118

6.4 Limitation 119

6.5 Future Work 120

6.6 Conclusion 121

**REFERENCES** 124

**APPENDICES** 134

## LIST OF TABLES


| | | Page |
|---|---|---|
| Table 2.1 | Requirement Boilerplate template | 18 |
| Table 2.2 | User Story vs Use Case Testing | 21 |
| Table 3.1 | Frequency of propose properties to be assessed in 93 gathered SRS | 32 |
| Table 3.2 | Sample assessment of FR | 40 |
| Table 3.3 | Hardware minimal requirement | 45 |
| Table 4.1 | Sample code sanitization structural assessment | 55 |
| Table 4.2 | Example pattern-matching denotation result | 56 |
| Table 4.3 | Sample code sanitization functional requirement assessment | 57 |
| Table 4.4 | Example of sentence sanitization | 59 |
| Table 5.1 | List of the functional requirement | 86 |
| Table 5.2 | Actual result refinement functional requirement | 87 |
| Table 5.3 | Possible similar functional requirement based on percentage | 90 |
| Table 5.4 | Result comparison and tokenization of *[F5]* and *[F8]* | 91 |
| Table 5.5 | Result comparison and tokenization of *[F7]* and *[F1]* | 92 |
| Table 5.6 | Present of datatype in functional requirement | 96 |
| Table 5.7 | Summary of structural measurement | 100 |
| Table 5.8 | Summary of consistency properties measurement | 102 |
| Table 5.9 | Summary of correctness properties measurement | 103 |
| Table 5.10 | Summary of preciseness properties measurement | 105 |

## LIST OF FIGURES


| | | Page |
|---|---|---|
| Figure 2.1 | Mental model of SRS document | 14 |
| Figure 2.2 | IEEE 830 standard table of content | 15 |
| Figure 3.1 | Research Question | 29 |
| Figure 3.2 | Frequency present of IEEE 830 topic | 31 |
| Figure 3.3 | Quality framework to measure SRS document | 34 |
| Figure 3.4 | Component Diagram | 46 |
| Figure 4.1 | Prototype activity diagram | 51 |
| Figure 4.2 | Sanitization of software requirement specification document | 54 |
| Figure 4.3 | Sample of pattern-matching code for structural assessment | 55 |
| Figure 4.4 | Sanitization of Functional Requirement Document | 57 |
| Figure 4.5 | Sample code to identify stakeholder | 58 |
| Figure 4.6 | Sample code to identify verb in sentence | 58 |
| Figure 4.7 | Sample code for similarity checking | 60 |
| Figure 4.8 | Sample code to detect present of stakeholder | 61 |
| Figure 4.9 | Sample code to detect present of possible design | 61 |
| Figure 4.10 | Sample code to check the present of vague word and datatype | 62 |
| Figure 4.11 | Entity relationship diagram | 63 |
| Figure 4.12 | Sample of add new project | 65 |
| Figure 4.13 | Sample portion SRS document contain ToC | 65 |
| Figure 4.14 | Sample result of text sanitization | 66 |
| Figure 4.15 | Sample result pattern-matching for structural assessment | 66 |
| Figure 4.16 | Graph presentation of detected topic | 67 |
| Figure 4.17 | Sample suggested synonym topic to IEEE 830 topic | 67 |

| | | |
|---|---|---|
| Figure 4.18 | Sample of text in FR document | 69 |
| Figure 4.19 | Sample on how to add stakeholder to the prototype | 69 |
| Figure 4.20 | Sample result FR refinement with similarity check | 70 |
| Figure 4.21 | Sample result consistency check | 71 |
| Figure 4.22 | Sample contain of Consistent Function page | 71 |
| Figure 4.23 | Sample contain of Inconsistent Function page | 72 |
| Figure 4.24 | Sample of test case | 73 |
| Figure 4.25 | Sample input in test case | 74 |
| Figure 4.26 | Sample of correctness assessment result | 75 |
| Figure 4.27 | Sample of preciseness assessment result | 77 |
| Figure 4.28 | Sample of datatype identification result | 77 |
| Figure 4.29 | Star graph represent the overall quality of assessed SRS document | 80 |
| Figure 5.1 | Sample of structural assessment and its result | 83 |
| Figure 5.2 | Graph preview structural assessment result | 85 |
| Figure 5.3 | Correct test case | 94 |
| Figure 5.4 | Incorrect test case | 95 |
| Figure 5.5 | Graph result of structural assessment | 99 |
| Figure 5.6 | Graph summarization quality properties assessment for first case study | 107 |
| Figure 5.7 | Graph summarization quality properties assessment for second case study | 108 |
| Figure 5.8 | Summary degree of each quality assessment for first case study | 109 |
| Figure 5.9 | Summary degree of each quality assessment for second case study | 110 |

## LIST OF ABBREVIATIONS

| | |
|---|---|
| CRUD | Create, Read, Update and Delete |
| FR | Functional Requirement |
| NFR | Non-Functional Requirement |
| RE | Requirement Engineering |
| SDLC | Software Development Life Cycle |
| SRS | Software Requirement Specification |
| ToC | Table of Content |

# CHAPTER 1

# INTRODUCTION

## 1.1 Background

Software engineering is a software process, which commonly consists of four fundamental activities. Those fundamental activities are software specification, software development, software validation and software evolution. This research concerns the RE process in software specifications.

RE process is including four phases which are feasibility study, requirement elicitation and analysis, requirement specification and requirement validation (Sommerville, 2016). The feasibility study focuses on the usefulness of the proposed system towards the organisation. The study included the system contribution towards the organisation, the capability to engineer the proposed system within budget and the capability of the proposed system to integrate with the others. The activities of requirement gathering, requirement classification and organisation, prioritization and negotiation and requirement documentation are the requirements of elicitation and the analysis. The requirement specification is the complete specified of the user and system specification after the requirement gathering process. While the requirement validation consists of the requirement checking process either it met the client's needs.

SRS is a documentation used to describe the user requirement and a detailed specification of the developed system. All aspects of the details must be specified before the project commerce (Pressman & Maxim, 2015). All of the detail mentions is refer to the produced SRS document. The SRS document should compose at least FR, NFR, level

prioritization of FR, target stakeholder, interface design and database design. A complete structure of the SRS document can be seen based on the topic suggested by IEEE 830 standard. All of the requirements were gathered through requirement gathering technique and they were divided into user and system requirements. Both of those requirements will be documented in the SRS.

This research concerns on the measurement of the SRS document quality. There are four qualities to be assessed; correctness, completeness, consistency and preciseness. The qualities were applied efficiently to measure the quality of SRS.

## 1.2 Problem Statement

Requirements gathered from the client were in a raw state and they were needed to undergo requirement elicitation process. RE is a practice of where the raw requirement is converted into useable FR. According to Zafar et al. (2018), neglecting the RE practice may impact the software development process. A survey done by Zafar et al. (2018) indicated that the main factors which can cause neglection of RE practices are limited budget, time, dedicated team and communication.

Besrour et al. (2016) created a survey to identify the challenges in the context of RE practices. The survey highlights one of the challenges which is undocumented FR and NFR, which can affect poor communication either between the developer team itself or between developer and client. Other than that, the survey is very crucial in order to collect vague and ambiguous requirements and poorly defined specifications. The usage of unstructured NL in defining the FR can cause ambiguity and poor understanding of the complex requirement.

Saito et al. (2014) and Takoshima and Aoyama (2015) highlight the issue regarding the structure of the SRS document. Those researchers arise the issue on how the FR and NFR are organized in the SRS document. Included also is the issue regarding the incapability of the current IEEE 830 standard topic to coop with heterogeneous domain in software development. However, they did express that the current IEEE 830 standard topic must be complied in order to organize the written FR and NFR.

There are two main problems that were highlighted in this research; unstructured structural of SRS and unstructured FR.

SRS document contains a number of elements as stated in IEEE 830 (IEEE Std. 830-1998, 1998). IEEE 830 standard had suggested a document structure that should be complied by the developer in order to produce a proper SRS document. The structure of the document shows the availability inside the content. Each SRS document structure may be different based on its domain study. The produced SRS document is based on the developer's understanding of the study domain. Due to no common standardisation which is compulsory to be followed, most of the documents are ill-defined, which can affect their quality.

Defining the FR is one of the main objectives of producing the SRS document. In order to write the FR, a clear specification needs to be defined. Clear specification resulted in consistency, correctness and preciseness. Commonly, FR is defined in NL. Unstructured FR is prone to inconsistency, ambiguity and incorrectness. NL is the most ambiguous model. There are numerous types of ambiguity (e.g., lexical ambiguity, syntactic ambiguity, semantic ambiguity, pragmatic ambiguity, vagueness and generality, language error ambiguity). Ambiguity causes multiple interpretations between the writer and reader of the document. This situation may affect the preciseness of the specified FR.

## 1.3 Objectives

The objectives of this research project are:

i. To identify the properties of the SRS document that affected the quality of the document. The properties are attribute of the quality which is needed in order to be assessed the SRS document structure and its FR. Assessment of this document's property is expected to give impacts to the quality of completeness, consistency, correctness and preciseness.

ii. To propose rules and equations based on the quality properties to be assessed. Each of the assessed quality has its own proposed properties. The completeness quality properties are proposed to assess the structure of the SRS document meanwhile the others are the FR. The rules and equations are proposed to comply with the software quality metric measurement for a quantitative study.

iii. To evaluate the proposed framework based on the software artefacts (e.g., SRS document). Based on the rules and equations, the quality framework was formulated. The framework is converted into a prototype. Train data as well as case study were used to evaluate the framework.

## 1.4 Hypothesis

The hypothesis for this research project are:

i. If the topic suggested in the IEEE 830 standard is followed, then the structure of the SRS document is complete and readable. The increment level of complexity in the software process can give an impact on the time for

developing the software product, meanwhile, the differences in the software production domain can cause the IEEE 830 standard to be ignored. IEEE 830 standard has been suggested as a standard guideline, which must be followed by any developer in order to generate the SRS document. As the SRS document is commonly written in NL, it is the developer's freedom to focus on their topic and ignore the standard. An automated document structure evaluation is proposed in order to ensure the developer follows the standard guideline.

ii. If the writing follows the RB template for the FR, it can reduce the ambiguity, inconsistency and incorrect conditions. One of the impacts of using NL in writing the FR is varied based on the interpretation of one individual to another. RB template restricts the usage of the word which can reduce the possible number of interpretations. It is also suggested that evaluation of the refined FR by proposing quality properties, can increase the FR quality.

## 1.5 Scope of Project

The scopes of the study can be simplified as below:

i. To assess the completeness structure of the SRS document based on the IEEE 830 topic as a standard guideline. 25 topics from IEEE 830 standard are extracted as a corpus to assess the document.

ii. To assess the structure of FR by restructuring it based on the RB template for the balance of its consistency, correctness and preciseness. There are some quality properties for consistency, correctness and preciseness which are

proposed and they are consisted of general properties. The proposed general properties are applied to assess the quality SRS document regardless of its domain.

iii. To develop the framework based on the assessed quality included its properties. The framework then converted into a prototype. The prototype is developed in a web-based environment. The prototype will only accept .docx format document as an input.

## 1.6 Thesis Structure

This thesis consists of six chapters. The first chapter is the introduction of the study, which consists of the background study, current issue with SRS document, objectives of the research, methodology used, scopes of study and the thesis structure.

The second chapter is the literature review, as it is very important in the study on the quality involve in SRS. The method used in assessing the structural and FR were extracted and evaluated. The chapter also includes the challenges of the current approach and the method which was adopted into the study.

The third chapter focuses on the proposed framework. The quality framework was proposed and discussed on the flow of the prototype, starting with the input until the production of results. The rules and equations for each quality were also proposed in this section.

The fourth chapter is about the implementation of the proposed framework. This section focuses on prototype architecture and workflow.

The fifth chapter focuses on the discussion. Two case studies were included in the prototype. The first case study is a case-control study. The input for the case-control study is predefine. The second case study involved the actual case study gather from the web site. The result from both case studies is clearly explained in the discussion section.

The last chapter is the conclusion. This conclusion section focuses on the achievement, contribution, limitation and future works.

## 1.7 Significant of Project

Throughout the project, there are numbers of benefits that can be expected, as stated as the following:

i. **Minimize human interference in the quality evaluation of the SRS document.** The evaluation of the structure is fully automated while as for FR, human interference is needed in certain areas.

ii. **Study on the relationship between the properties inherits by each quality assessed.** There was a study done on ISO 25010, which showed mutual influences between each of the quality assessed. Those qualities were supported by the properties proposed toward them as such as performance quality which owned the properties of time, resources and capacity. This is an advantage for the researcher to evaluate the produced SRS document based on the properties stated.

iii. **Increase client and developer acceptance.** This study only involves the developer side which means there is no involvement of clients within the assessed project. The idea is to increase the degree of acceptance from the

clients. The produced SRS document must achieve a certain degree of quality before presenting it in front of them. By using the prototype, the SRS document is evaluated into a certain degree of quality before it is submitted to the client. The result from the evaluation shows the maturity of the produced SRS document as well as an understanding of the developer toward the written FR.

iv. **Minimal time constraint**. Triple constraints; time, budget and quality are the main issues when developing a system. This prototype can minimize the time needed in order to understand the complexity of the written FR.

v. **Promote good RE practice.** RE practices are included in a crucial phase in software development. Good RE practices can produce mature documentation of FR and it is important as the quality of software production depends on it. This prototype promotes the completeness quality of the document structure. The quality of FR can refine unstructured FR into structured FR by adopting the RB template. Structured FR resulted in the quality of consistency, correctness and preciseness.

## 1.8 Summary

Thus, this chapter introduced the area of study which was the assessment on the quality of the SRS document. The current problem with assessing the SRS document has been identified which is the lack of standardization is the main issue. The objectives and methodology of the research are already defined. The scope of the study focuses on two elements; structural and FR. There are four qualities that were applied to assess the SRS document which is completeness, consistency, correctness and preciseness.